CHAPTER THREE

False Recollection: Empirical Findings and Their Theoretical Implications

Jason Arndt

Contents

1. Introduction 82
2. False Memory Due to Familiarity and Recollection 83
3. Behavioral Evidence That People Experience Recollective Memory Errors 85
   3.1 The Misinformation Effect 87
   3.2 Imagination Inflation 90
   3.3 List-Learning Paradigms 93
   3.4 Summary of Behavioral Evidence and Implications 98
4. Brain-Based Evidence for False Recollection 99
   4.1 ERP Studies 100
   4.2 FMRI and PET Studies 102
   4.3 Summary of Brain-Based Evidence and Implications 103
5. Theories of False Recollection 104
   5.1 Fuzzy-Trace Theory 105
   5.2 Activation-Monitoring Theory 108
   5.3 Global-Matching Models 112
   5.4 Comparison of Theories 117
6. Future Considerations 118
References 119

Abstract

This chapter reviews the literature documenting particularly compelling false memories—memory errors that people subjectively experience as similar to authentic memories (referred to as *false recollections*). Behavioral and neuroimaging evidence examining three major phenomena that produce false recollection (the misinformation effect, imagination inflation, and DRM lure errors) are reviewed, with the general goal of understanding the kinds of memorial information that underlies them. In particular, this chapter was motivated by consideration of whether familiarity-based processes are sufficient to explain false recollection, or if

Psychology of Learning and Motivation, Volume 56
ISSN 0079-7421, DOI 10.1016/B978-0-12-394393-4.00001-7
© 2012 Elsevier Inc.
All rights reserved.

evidence exists that false recollections are caused by processes similar to those mediating authentic recollection, such as the retrieval of encoding context. Following the review of evidence documenting false recollection, three major theoretical frameworks that explain false recollection are evaluated in light of the evidence, with the goal of advancing understanding of the memory representations and retrieval processes that underlie false recollection.

1. Introduction

Human memory is subject to a multitude of errors, including, but not limited to, misattributions of familiarity (Jacoby & Whitehouse, 1989), confusions between reality and imagination (Johnson & Raye, 1981), susceptibility to misinformation (McCloskey & Zaragoza, 1985), eyewitness identification errors (Wells, Rydell, & Seelau, 1993), and beliefs that we have encountered events that were never experienced (Deese, 1959; Roediger & McDermott, 1995). While understanding these effects and their bases provides a useful basis for constructing recommendations for applied practice (e.g., Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000), they are also a useful basis for understanding the operating principles of human memory (Kimball, Smith, & Kahana, 2007).

The central issue this chapter will address is why compelling memory errors—those that people strongly believe are authentic memories—arise. Thus, rather than errors that occur because people believe events are familiar (Diana, Peterson, & Reder, 2004), the cases this chapter will consider are those errors that are accompanied by particularly strong convictions, leading people to claim to be able to recollect details of events they did not experience (Roediger & McDermott, 1995) and to believe events they did not experience were encountered with high confidence (Anastasi, Rhodes, & Burns, 2000); that is, this chapter is concerned not with false memories in general, but with false *recollections* of prior events (see Brainerd & Reyna, 2005; Gallo, 2006, 2010 for general reviews of false memory). While such false recollections have been the subject of study and theorizing in the past (Arndt, 2010; Brainerd, Wright, Reyna, & Mojardin, 2001; McCabe & Geraci, 2009), this chapter will review several major sources of evidence that document the nature of these compelling memory errors, and will use that evidence to evaluate the mental processes that underlie mistaken recollections. I begin by reviewing two general explanations of compelling memory errors that suggest very different interpretations of false recollection: that they represent erroneous attributions of familiarity,

and that they represent retrieval of encoding context. Second, I review findings in the literature that have been purported to demonstrate false recollection, with the goal of arguing that at least some cases of false recollection represent instances where people are retrieving evidence from memory that events they did not experience were nevertheless encountered in a specific encoding context. Third, I consider several major theoretical explanations of false recollections, and evaluate those theories in light of the data reviewed here. Finally, I close with a consideration of future questions that may be useful for research examining false recollection to pursue.

2. False Memory Due to Familiarity and Recollection

Extensive evidence suggests that recognition memory can be best understood by considering the operation of at least two different memory processes, familiarity and recollection (Diana, Reder, Arndt, & Park, 2006). Familiarity is a memory process that provides a nonspecific sense for an event's prior occurrence, but generally does not provide people with more detailed information about a previous experience, such as the context in which a memory was experienced (Yonelinas, 2002). Recollection, on the other hand, is a memory process that is underlain by retrieval of the context in which an experience occurred (referred to *encoding context* hereafter), such as where and when it occurred, as well as other events that occurred close in proximity to it (Yonelinas, 2002). Traditionally, recognition memory false alarms have been assumed to occur because of the influence of familiarity (Yonelinas, 2002), likely because the items that produce false memory are, by definition, not experienced in the encoding context people are asked to make memory judgments about. Thus, it seems conceptually implausible for false memories to involve retrieval of encoding context.

There are several key challenges to the general notion that false memories are underlain by familiarity. For example, people sometimes believe they can recollect the occurrence of events they did not experience (Roediger & McDermott, 1995). Similarly, people will claim they actually performed actions they only imagined performing (Goff & Roediger, 1998). In addition, people are willing to attribute misinformation to specific circumstances other than the source of the misinformation, such as an event they witnessed (Zaragoza & Lane, 1994). Each of these basic findings suggests that people experience some memory errors in a way that mimics their experience of recollecting a previous occurrence, which seemingly creates a dilemma for theories attributing memory errors to familiarity-based processes. However, there exist theories that suggest

memory processes that produce a strong sense of familiarity, such as unstudied test items that resemble a strong gist memory trace (Brainerd et al., 2001), can cause people to the erroneously believe that they recollect false memories (referred to as *phantom recollection*). Similarly, theories of recollection phenomenology that are based in signal-detection theory (Green & Swets, 1966) suggest that strong familiarity can underlie the belief one can recollect a previous occurrence (Donaldson, 1996; Hirshman & Master, 1997). Thus, although people's subjective experience suggests they can recollect (1) misinformation as part of an event they witnessed, (2) performing actions they only imagined, and (3) events that never occurred, those subjective experiences can be explained using only familiarity-based mechanisms. While familiarity-based mechanisms can account for people's subjective experience of false recollection, an alternative explanation for why people find some memory errors highly convincing is that these memory errors are supported by recollection-based processes (Arndt, 2010; McCabe & Geraci, 2009). This general claim argues that events that were not experienced can, in some cases, produce retrieval of the encoding context of an event that was actually experienced, which is mistakenly used to validate a person's false memory experience. Thus, there are two general explanations for why people experience some false memories as particularly compelling, and those two general explanations differ greatly in their claims about the memorial information that underlies false recollections.

The central question this chapter will seek to resolve is the extent to which people's experience of false recollection is underlain by familiarity-based and recollection-based processes. Given that the simpler explanation is that false recollection is subserved by familiarity-based processes, we must seek evidence that demonstrates that false memories cannot be explained by constructs often proposed to underlie familiarity, such as memory trace strength, in order to adequately justify the view that recollection-based processes contribute to people's belief they can recollect false memories. Further, empirical support for the hypothesis that false recollection involves retrieval of recollective information requires evidence that not only shows that familiarity-based explanations are insufficient to fully explain false recollection, but also supports the notion that false memories can cue retrieval of specific details that were part of the encoding context of authentically experienced events—a core claim of the view that recollection underlies some false memories.

The review that follows uses this general theoretical framing to organize each section describing the major phenomena that have been argued to produce false recollection. First, major findings within an area of the literature are reviewed that document the strong subjective experience of recollection that often accompanies false memories. Second, the ability of familiarity-based memory mechanisms to explain the subjective

experience of false recollection is described, as well as ancillary evidence that supports familiarity-based explanations of false recollection. Finally, to the extent the evidence exists, empirical findings that demonstrate that false memories cue retrieval of objectively verifiable details from a specific encoding context will be described to document the phenomena for which recollection-based processes contribute to false memories. Consequently, this chapter will distinguish between empirical findings that can be explained fully with familiarity-based processes, and those that cannot be explained without the contribution of recollection-based processes.

3. Behavioral Evidence That People Experience Recollective Memory Errors

There are several basic memory paradigms that have been shown to produce particularly compelling memory errors, such that people believe they can recollect specific details of an event they did not experience. These errors arise in paradigms investigating the misinformation effect (Loftus, 1975; Meade & Roediger, 2002), confusions between imagination and performing actions (Goff & Roediger, 1998; Seamon, Philbin, & Harrison, 2006; Thomas & Loftus, 2002) or perceiving objects (collectively referred to as *imagination inflation* hereafter; Henkel, Johnson, & De Leonardis, 1998; Lyle & Johnson, 2006), and list-learning paradigms (Deese, 1959; Roediger & McDermott, 1995). Further, parallel lines of research using neuroimaging have documented the brain processes associated with memory errors, which can also provide insight into the memorial information being accessed when memory errors are committed. I will review the basic findings in each of these lines of work in order to evaluate how strongly they suggest that the memory errors they produce are truly false recollections, beginning with the major results from studies of behavioral outcomes. Following the review of behavioral outcomes, I will review results from studies of memory errors using neuroimaging techniques, again with the goal of examining how strongly they suggest memory errors produce retrieval of encoding context. As a supplement to the review that follows, Table 1 presents the general empirical trends within the misinformation effect, imagination inflation, and list-learning literatures. Table 1 is organized in parallel with the major categories of evidence reviewed in this chapter. The first column summarizes the existence of evidence that people find false memories subjectively compelling, the second column summarizes the existence of behavioral evidence that false memories produce retrieval of encoding context, and the third (event-related potentials; ERP) and fourth (positron emission

Table 1 Summary of Outcomes from Research Investigating False Recollection

Behavioral paradigm	Nature of evidence			
	Subjective belief	Retrieval of encoding context	ERP activation	PET/fMRI activation
Misinformation effect	*Evidence*? Yes *Example*: Misinformation attributed to original event.[a]	*Evidence*? No *Example*: Context similarity between misinformation and original event does not increase misinformation effect.[d]	*Evidence*? Unknown	*Evidence*? Yes *Example*: Misinformation produces MTL activation at retrieval.[h]
Imagination inflation	*Evidence*? Yes *Example*: Imagined items claimed to be performed.[b]	*Evidence*? Yes *Example*: Imagined items attributed to source of similar perceived item.[e]	*Evidence*? Unknown	*Evidence*? Unknown
List learning (DRM)	*Evidence*? Yes *Example*: Remember judgments given to lure items.[c]	*Evidence*? Yes *Example*: Lures attributed to source of strong associates more often than source of weak associates.[f]	*Evidence*? No *Example*: Lure errors do not produce late left parietal ERPs.[g]	*Evidence*? Yes *Example*: Lure errors produce MTL activation during retrieval.[i]

Note: Superscripts indicate citations for example evidence.

[a] Zaragoza and Lane (1994).
[b] Goff and Roediger (1998).
[c] Roediger and McDermott (1995).
[d] Mitchell and Zaragoza (2001).
[e] Lyle and Johnson (2006).
[f] Hicks and Hancock (2002).
[g] Curran et al., 2001.
[h] Stark et al., 2010.
[i] Cabeza et al., 2001.

tomography (PET)/functional magnetic resonance imaging (fMRI)) columns summarize the evidence from neuroimaging studies documenting that false memories activate brain areas associated with retrieval of encoding context. Further, where evidence exists, Table 1 provides an example finding and representative citation that supports (*Evidence*? *Yes* entries) or does not support (*Evidence*? *No* entries) the hypothesis that false memories produce each source of evidence for false recollection reviewed in this chapter.

3.1. The Misinformation Effect

Research on the misinformation effect has a long and rich history in the human memory literature, beginning with the pioneering work of Loftus (1975; Loftus & Palmer, 1974). In a typical misinformation effect study, participants witness an event, usually visually via slides (Loftus & Palmer, 1974) or a video (Chan, Thomas, & Bulevich, 2009). At some point after witnessing the event, but prior to having their memory tested, participants are provided with misleading information about the event they witnessed, via a narrative (Chan et al., 2009; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001) or misleading questions (Zaragoza & Lane, 1994). Finally, after the provision of misinformation, participants' memory is tested, with results typically showing that their memory for the events they witnessed was influenced negatively by the misinformation, such that they believe misinformation was actually seen during the slides or video that they observed. Thus, misinformation effects document cases where people confuse a verbal description of an event with a previous visual experience.

There are several findings that support the hypothesis that the provision of misinformation can give rise to a strong belief that misinformation was experienced as part of the same experience as the witnessed event (i.e., recollected). Most of this evidence comes from research examining source attributions that participants make following their exposure to misinformation. In particular, some studies have asked participants to evaluate if they remember experiencing items as part of an event they witnessed (Zaragoza & Lane, 1994). In contrast to simply expressing the general belief that they experienced misinformation, source judgments provide a metric of the specificity of a person's beliefs about where or when they experienced misinformation (Johnson, Hashtroudi, & Lindsay, 1993). Typical studies using source memory to tap the bases of peoples' beliefs about misinformation provide subjects with the option of indicating that they believe misinformation was experienced only during the witnessing of the original event, was experienced only during the provision of postevent information, was experienced during both, or was not experienced at all. The important data from such studies is the likelihood

people claim to have experienced misinformation during the witnessing of the original event, as indexed by the probability they believe the misinformation was only seen during the witnessed event added to the probability they correctly believe that the misinformation occurred during the postevent information, but also believe it was encountered during the witnessed event (Zaragoza & Lane, 1994). Generally, studies using such source judgments have shown that people will attribute misinformation arising from a postevent narrative or postevent questions (Zaragoza & Lane, 1994) to a witnessed event. Similar evidence has been found when a recall partner serves as a source of misinformation about the contents of studied pictures (Meade & Roediger, 2002). Thus, these findings serve as preliminary evidence that people believe they are retrieving authentic visual memories, even though the source of those memories was a verbal narrative, misleading question, or a recall partner.

There are two additional findings from misinformation effect studies that underscore the strong belief people have that their memory for misinformation was part of the originally encoded event. First, while explicit warnings that they may have received misinformation decreases the rate at which they mistakenly claim misinformation was part of the witnessed event, people nevertheless continue to attribute misinformation to the witnessed event at a rate that is higher than for control items (Meade & Roediger, 2002). Second, when asked to rate their confidence that misinformation was experienced during a witnessed event, people tend to show high confidence in their mistaken source judgments (Zaragoza & Lane, 1994), suggesting such source misattributions are not simply guessing errors, but instead are based upon strongly held convictions. Thus, evidence suggests that people strongly believe that misinformation was authentically experienced as part of a witnessed event, which is consistent with the view that they experience an authentic, albeit mistaken, recollection of the misinformation occurring as part of the originally witnessed event.

However, because all of the above findings document cases where people provided subjective judgments about the contents of their memories, and none of the findings document cases where misinformation produced retrieval of specific details from the originally experienced events, it is possible for a familiarity-based explanation to account for the data. A familiarity-based theory would claim that some misinformation produces strong memories, which leads people to believe the misinformation must have been experienced visually, causing them to attribute misinformation to the originally witnessed event. Importantly, this familiarity-based explanation is also consistent with findings showing that misinformation that is attributed to the witnessed event tends to be endorsed with high confidence (Zaragoza & Lane, 1994), because highly familiar misinformation would also tend to be judged with high confidence.

Two additional results support the notion that misinformation effects can be explained by familiarity-based processes. First, the more often people are exposed to misinformation, the more likely they are to attribute the misinformation to the originally witnessed event (Meade & Roediger, 2002; Mitchell & Zaragoza, 2001; Zaragoza & Mitchell, 1996). While it is possible that repeated exposure to misinformation would enhance the likelihood it would be integrated with encoding context from a witnessed event, it is also the case that the more often misinformation is encountered as part of a postevent narrative (Mitchell & Zaragoza, 2001; Zaragoza & Mitchell, 1996), or as part of a postevent recall protocol (Meade & Roediger, 2002), the more likely a person would be able to accurately recollect that the misinformation was part of a postevent narrative or recall protocol, causing misinformation effects to decline with repetition. In contrast, a familiarity-based view has a straightforward explanation for why repetition of misinformation increases misinformation effects—the more often misinformation is encountered, the more familiar it is, and thus the more likely that familiarity will be attributed to the originally witnessed event. A second result that favors a familiarity-based explanation of misinformation effects is that increasing contextual overlap between a witnessed event and misinformation does not increase attributions of misinformation to the witnessed event, and erroneous recall of misinformation as part of a witnessed event decreases (Mitchell & Zaragoza, 2001). The fact that contextual overlap between a witnessed event and misinformation does not influence the misinformation effect is consistent with the notion that the misinformation effect is familiarity driven, given that familiarity is generally viewed as providing an acontextual sense of an event's prior occurrence based upon memory trace strength. The latter finding that greater contextual overlap actually reduces the likelihood that misinformation is recalled as part of the originally witnessed event is not immediately explicable by familiarity-based views of misinformation effects, but is also an outcome that is the opposite of what a recollection-based explanation of false recollection would expect. Thus, this data pattern is more appropriately viewed as inconsistent with the recollection hypothesis than support for misinformation effects being driven by familiarity. In summary, while the provision of misinformation about a witnessed event often can produce subjectively convincing false memories, such that people believe they can recollect misinformation as having occurred as part of a witnessed event, and do so with high confidence, the existing evidence from misinformation effect studies does not document the retrieval of authentic encoding context when misinformation is falsely remembered (see Table 1, first row). As a consequence, evidence from behavioral investigations of the misinformation effect can be explained fully with familiarity-based processes and do not provide evidence that recollection contributes to people's experience of false recollection.

3.2. Imagination Inflation

A second line of evidence that has been suggested to demonstrate false recollection examines imagination inflation (Goff & Roediger, 1998). In the first phase of a typical imagination inflation study, participants listen to a variety of action statements (e.g., break a toothpick in half), and are asked to either perform or imagine performing those actions. In a second phase of the study, participants are simply asked to imagine performing some of the actions they performed or imagined during the first phase, as well as a set of new actions that they did not encounter during the first phase. Finally, on a later memory test, they are asked to judge what they did during the first phase of the study—did they hear about a particular action, and if so, was the action actually performed or was it simply imagined? The typical finding in such studies is that people believe that they performed actions that they simply imagined performing, both for items that were only imaged during the first phase of the study as well as for items that were never experienced during the first phase of the study, but were imaged during the second phase (Goff & Roediger, 1998; Seamon et al, 2006). Thus, the simple act of imagining an event occurred, even if the imagination occurred in a phase of the study where no actions were performed, could be confused with the experience of actually performing that action. The fact that participants believe they actually performed an action that they did not physically engage in has been taken as evidence they believe they are recollecting a prior experience (Goff & Roediger, 1998).

Four findings from the imagination inflation literature support the view that people believe they are recollecting the experience of having performed actions they only imagined. First, people claimed to be able to recollect vivid details of performing many of the actions that they only imagined, even though they were given the opportunity to indicate they were certain they had performed the action, but were unable to recollect any details of its occurrence (Thomas & Loftus, 2002). Second, people claim that they both performed and could recollect details of performing bizarre actions (e.g., kiss the magnifying glass) that they only imagined performing (Thomas & Loftus, 2002). Third, the more perceptual details a person imagined about actions they did not perform, the more likely they were to believe they performed the action, and to believe they could recollect having performed the action, regardless of whether it was a familiar or a bizarre action (Thomas, Bulevich, & Loftus, 2003). Fourth, and finally, imagination inflation for familiar and bizarre events occurred even when people performed or imagined events in contextually rich settings, such as a series of distinct places on a college campus (Seamon et al, 2006). Thus, people show strong conviction that their mistaken recollections of events they only imagined are authentic and

show that strong conviction for bizarre actions that are unlikely to have been performed previously in their life, suggesting they are not simply confusing imagining an event during the experiment with an extra-experimental episode where they had actually performed the imagined action. Further, providing people with a contextually rich environment that should be optimal for distinguishing performed from imagined actions failed to eliminate the imagination inflation effect (Seamon et al., 2006). Indeed, in Seamon et al's (2006) study, overall discrimination was exceedingly high, with approximately 85% of the bizarre actions people actually performed being accurately judged as performed, and only 2% of entirely novel actions (neither performed nor imagined during the study) being mistakenly judged as performed. Thus, despite overall excellent memory for actions that were performed, and very low levels of mistakes, people still believed that they had performed an action, whether familiar or bizarre, when they had only imagined it.

These four results combine to suggest that people find actions they have only imagined performing to be strikingly similar to events they have performed, which in some instances leads them to believe the imagined action was actually performed. As with misinformation effects, the conviction with which people believe actions they only imagined were actually performed is consistent with the notion that they are experiencing an authentic recollection that they performed those actions. Further, that participants hold this belief for bizarre actions, which they likely thought about for the first time during the experiment, is consistent with the notion that they are not simply confusing their imagination with an extra-experimental action they had performed in the past, but instead are experiencing authentic mistaken recollection of having performed something they only imagined.

Although these findings from imagination inflation research are consistent with the notion that people retrieve recollective information that supports false recollections of performing actions they only imagined, it is also possible for a familiarity-based view to account for each of these results, given that each finding documents a subjective state of recollection without providing evidence that people actually retrieved specific episodic details that occurred when actions were performed. Thus, as with outcomes from misinformation effect research, a familiarity-based view could claim that people experience high levels of confidence and memory strength for some of the actions they only imagined, which in turn leads them to believe that those actions were likely to have been performed. Further, the fact that imagery, a technique that creates strong memories (Bower & Winzenz, 1970), is used for actions that are later confused with performing those actions makes a familiarity-based view of imagination inflation errors particularly plausible. Finally, one additional result from

the imagination inflation literature seems to fit more naturally with a familiarity-based view than the view that imagination produces retrieval of encoding context, producing false recollection. Specifically, imagination inflation has been found to increase with repeated imaginings (Goff & Roediger, 1998). While this outcome can be explained by the view that the more a person imagines executing an action, the more likely they are to produce a memory that overlaps with the motor and perceptual experience of actually executing the action (Goff & Roediger, 1998), it also seems likely that increasing the number of times an action was imagined would help solidify their recollection that they imagined the action, which should in turn *decrease* the likelihood they would falsely recollect performing the action. In comparison, a familiarity-based explanation would expect an increase in familiarity for the imagined event, and thus increased imagination inflation, the more an action is imagined. Thus, as with misinformation effects, it is possible for familiarity-based processes to explain the imagination inflation effects discussed to this point, because each of the key findings was based upon participants' subjective belief that they performed an action that they had only imagined, and not on objective evidence that participants retrieved encoding context from a performed action when tested with an imagined item.

Subsequent work examining source confusions between imagined and perceived events is, however, inexplicable from the standpoint of a familiarity-based theory (Henkel & Franklin, 1998; Lyle & Johnson, 2006). In these studies, participants were shown a study list that consisted of two basic types of trials. On some trials, they viewed a picture of a simple object, while on other trials, they were asked to imagine what an object looked like. Embedded within the series of pictures people viewed and the objects people imagined were item pairs constructed such that some of the imagined objects were similar in appearance to objects that were viewed as pictures, some of the imagined objects were conceptually related to objects that were viewed as pictures, and some of the imagined objects were neither perceptually nor conceptually related to any of the pictures (control items). On a later memory test, people were asked to judge whether each test item was viewed as a picture, was imagined, or was not encountered. The typical finding from research using this paradigm is that people tended to believe imagined items were viewed as pictures more often when the imagined item was perceptually or conceptually related to an item that was seen as a picture during study (Henkel & Franklin, 1998).

Although this basic finding can be explained by a familiarity-based view of false recollection, subsequent research has documented key evidence supporting the notion that people are drawing on specific episodic memory representations to make these judgments. Specifically, when asked to judge the exact location or color of an item they only imagined

during encoding, participants chose the location or color of the item that was perceptually or conceptually similar to the imagined item at levels that exceeded chance (Lyle & Johnson, 2006). The systematicity with which people attributed imagined items to the source of pictures that were conceptually or perceptually related to the imagined item is critical for justifying the view that these imaginations represent authentic cases of false recollection. Specifically, an explanation that argues source misattributions for imagined items are simply due to high levels of familiarity can explain a general tendency to believe that an imagined item was perceived or performed, but does not have a mechanism to explain why people systematically chose a particular place or color in which the imagined item was experienced. Rather, a familiarity-based explanation would have to argue that source misattributions are guesses, because familiarity processes produce a signal that indicates the strength of evidence in memory that an imagined item was encountered, but not information about where or in what format it was encountered. Thus, these source judgments for imagined events are the first data pattern that conclusively supports the notion that people's experience of false recollection can be based upon retrieval of encoding context from memory, and thus recollection-based processes.

Provided the results of Lyle and Johnson (2006) reflect the memory processes that typically produce imagination inflation, these results suggest that when people retrieve memories of events they only imagined, they sometimes access source information about events that were actually perceived or performed during encoding, which in turn leads them to believe that an event they only imagined was not only perceived, but perceived in a particular context (e.g., location or color). That is, imagination not only produces the subjective experience of perceiving an object or performing an action, but also produces memory representations that erroneously enable retrieval of recollective information, such as the contextual information that is encoded and stored along with perceived objects or performed actions. In sum, although most of the results from the imagination inflation literature are based upon participants' subjective conviction that an imagined item was perceived or performed, and thus can be explained using only familiarity-based mechanisms, the results from a few studies of imagination inflation support the hypothesis that retrieval of encoding context contributes to false recollection, which in turn underlies peoples' beliefs that their false memories are authentic recollections (see Table 1, second row).

3.3. List-Learning Paradigms

As with misinformation effect and imagination inflation research, studies examining false recollections in list-learning paradigms have been claimed

to provide evidence for false recollection (Dewhurst, 2001). The vast majority of the research in list-learning paradigms that has been claimed to produce false recollection comes from semantic/associative memory errors in the Deese–Roediger–McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) paradigm. In this paradigm, participants are presented with a series of items (e.g., *chirp*, *sparrow*, *bluejay*, *canary*, *feathers*, *nest*, *pigeon*, and *robin*) that are all related to a single unstudied item, known as the *lure* (e.g., *bird*). Of critical importance in this paradigm are the errors people make to the lure item and the characteristics of peoples' beliefs about lure errors. Regarding error rates, lure items produce higher false alarm rates than other types of lures on recognition memory tests, such that they substantially exceed the false alarm rates of other lures that are not associated with studied items, regularly producing false alarm rates in excess of 50%, and sometimes producing false alarm rates that are similar to hit rates for studied items (Roediger & McDermott, 1995). Further, in free recall tasks, which rarely produce high rates of intrusions, participants regularly recall lure items as having occurred on the study list, suggesting they strongly believe the lure was encountered in the same episode as the study items (Deese, 1959; McDermott, 1996; Payne, Elie, Blackwell, & Neuschatz, 1996; Roediger & McDermott, 1995). Peoples' belief that they can recollect lures from the study list is supported by direct measures of the subjective states that accompany lure errors. People will regularly claim to be able to recollect lures' presentation on the study list as opposed to simply judging lures to be familiar (Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Roediger & McDermott, 1995). In addition, people generally have high confidence that a lure was studied even though lures produce recollection of fewer perceptual details than studied items (Mather et al., 1997; Norman & Schacter, 1997) and people are willing to attribute lures to a source that was encountered during encoding (Hicks & Hancock, 2002; Hicks & Starns, 2006a; Mather et al., 1997; Roediger, McDermott, Pisoni, & Gallo, 2004). Thus, people's subjective experience with lure items suggests that they believe they can recollect the occurrence of lures from a recent encoding episode.

One further line of evidence that people strongly believe lures were studied is that lure errors are strikingly robust in the face of manipulations intended to eradicate them (Gallo, Roberts, & Seamon, 1997; Israel & Schacter, 1997; McDermott & Roediger, 1998; Multhaup & Conner, 2002). For example, while explicit warnings against making errors tend to reduce lure error rates (McDermott & Roediger, 1998; Gallo et al., 1997), they remain robust and well above baseline error rates. Similarly, asking participants to carefully assess their memories for information that is indicative of having actually experienced an item during encoding, such as perceptual details, their thoughts about the word during encoding, and

feelings they had about the word during encoding reduces lure errors rates, although they remain well above baseline error rates (Mather et al., 1997). Further, asking participants to justify their experience of recollecting a lure item does not seem to strongly reduce the likelihood that people believe they can recollect lure items (Norman & Schacter, 1997), and guiding participants to consider a variety of memory characteristics that should enable them to realize lure errors were not encoded (e.g., what the word sounded like, where in the list it occurred, what words it was next to during encoding, what the person's reaction to the word was, etc.) failed to eradicate lure errors, leaving them well above baseline error rates (50% lure errors vs. 8% baseline errors; Norman & Schacter, 1997, Table 4, data for younger adults). Building off of these studies that suggest there are differences in the perceptual details that support accurate memories and lure errors, Israel and Schacter (1997; see also Schacter, Israel, & Racine, 1999) demonstrated that pictorial encoding of a lure's associates substantially decreased lure errors, although lure error rates again remained well above baseline (e.g., 35% lure errors vs. 8% baseline errors; Israel & Schacter, 1997, Experiment 2). Finally, one study combined a number of techniques designed to improve discrimination between studied items and lures by providing participants with an example DRM study list and its accompanying lure item prior to study, instructing participants to try to figure out what the lure was so they could avoid endorsing it as studied, testing memory immediately after each lure's associates were encoded, and offering participants the opportunity to attribute lures to a nonlist source (e.g., the participants' thoughts or that the lure tied together the study items, but was not actually studied; Multhaup & Conner, 2002). Although those procedures reduced the probability a lure was endorsed as studied compared to standard testing conditions, lure errors remained well above baseline error rates (~40–50% lure errors vs. 0–2% baseline errors; Multhaup & Conner, 2002). The fact that lure errors are robust in conditions where strict warnings are provided, participants' attention is drawn to the memorial details that should distinguish between studied items and lures, and study items are encoded as pictures suggest that the memorial information supporting lure errors is sufficiently compelling and people are unable to distinguish lures from authentic memories.

As with misinformation effects and imagination inflation, the fact that people find lure errors subjectively compelling does not sufficiently support the view that these errors are underlain by actual recollection. For example, although lure errors are exceedingly robust in the face of efforts to reduce them with warnings and other metamemorial techniques is consistent with the view that lures produce (erroneous) retrieval of encoding context, it nevertheless is possible that lures simply inspire very strong familiarity, which in turn explains their robustness to efforts designed to limit them. However, as with imagination inflation studies,

there exists evidence that strongly suggests encoding context is retrieved when a lure item is tested, which in turn supports the notion that lure errors can be underlain by actual recollection.

One line of evidence that supports the notion that lure items produce retrieval of encoding context comes from studies of source memory judgments to lure items. As noted above, people are quite willing to attribute lures to a specific studied source when given a source memory test (Anastasi et al., 2000; Hicks & Hancock, 2002; Hicks & Starns, 2006a; Mather et al., 1997; Roediger et al., 2004). Importantly, there are three key results from studies of lure source memory judgments that document that lures inspire retrieval of encoding context despite the fact that they were not actually experienced during encoding. First, when people were asked to make source judgments, the probability they believe lure items were studied *increased* relative to when they were asked to make old–new recognition judgments (Hicks & Marsh, 2001). This result suggests that when people were asked to focus on retrieval of contextual information while making source judgments, they were actually *more* error prone than when they completed a task, old–new recognition, that relies to a greater extent on familiarity. Second, when a lure's studied associates are presented in a single source, participants disproportionately choose that source when they endorse the lure as studied (Hicks & Hancock, 2002; Mather et al., 1997; Roediger et al., 2004). This result suggests that people are not simply guessing when they indicate a lure was studied in a specific source, but instead are relying on specific retrieved information from memory to judge a lures' source, even though lures were not experienced during encoding. Third, and most important, when a lure's associates were presented in two different sources (e.g., a male and a female voice), people tend to attribute the lure to the source that is correlated with its stronger associates (Hicks & Hancock, 2002; Hicks & Starns, 2006a). This final result is particularly important because it suggests that participants access highly specific memorial information about the sources that were used to encode a lure's associates when their memory is probed with a lure item. Specifically, they access information about the source that lures are most strongly related to, and more often retrieve that source information when they make a judgment about a lure's source even though the lure was never actually experienced during encoding. Taken together, these three key results from studies of lure source attributions support the hypothesis that lure items produce retrieval of the specific encoding contexts in which their studied associates were encountered.

A second line of research examining contextual reinstatement between encoding and retrieval complements the evidence from studies of source memory for lure items. It is well known that memory benefits from contextual reinstatement, a finding that has been argued to indicate that memory representations are associated with encoding context

(Smith, 1979). Further, Reder, Donavos, & Erickson, (2002) demonstrated that using relatively unique contexts (unusual-looking fonts) produced context-reinstatement effects in recognition memory, which suggests that recognition memory also involves retrieval of encoding context. This basic finding has been used to examine how lure item errors are impacted by contextual overlap between encoding and retrieval. In a typical experiment, the study items related to a lure were presented in unusual-looking fonts, with the font used to present a specific lure's associates differing across sets of associates that were related to different lures. This design feature enabled evaluation of the extent to which testing lures in a visual context that was present at a specific point in time during encoding—the point when that lure's associates were studied—influenced lure errors (Arndt, 2010). At test, lures were either presented in the font that was used to encode its associates (the *match* condition) or were presented in a font that was studied, but was used to encode a different lure's associates (the *mismatch* condition). Given that contextual overlap enhances recognition memory, one might anticipate that testing lures in a font that was used to study its associates would reduce lure errors, because the lure's test cue would include perceptual information (font) that (1) does not match the lure's corresponding representation in memory, and (2) provides a cue to retrieve the lure's studied associates, which would enable the lure to be rejected as not studied. In contrast to this expectation, lure errors are generally *higher* when they are tested in a font used to study their associates compared to a font that was not used to study their associates (Arndt, 2006, 2010). Importantly, because lures in the mismatch condition were tested in a font that was studied, but was not used to present the tested lure's associates during encoding, explanations that highlight the role of perceptual fluency (Jacoby & Whitehouse, 1989) or general increases in familiarity are not tenable; this comparison instead suggests that people are accessing specific episodic information that only occurred during encoding of the lure's associates. Thus, as with recognition memory for studied items, lure errors increased when they were tested in a context that was encountered during encoding, suggesting the lure item cued retrieval of the context that was present when its associates were encoded.

In addition to the basic result documenting that lure errors are more common when a retrieval cue reinstates the encoding context of the lure's associates, two results from studies of how context influences lure errors support the notion that people are retrieving specific episodic details when they endorse lure items as studied. First, when a lure's associates were studied in two different fonts such that one font presented the lure's strong associates and the other font presented the lure's weak associates, people made more errors when the lure was tested in the font used to present its strong associates than when it was tested in the font used

to present its weak associates (Arndt, 2006). As was the case with source attributions varying with associative strength (Hicks & Hancock, 2002; Hicks & Starns, 2006a), this result suggests that people access highly detailed contextual information when their memory is tested with a lure item. Second, when a lure's associates were studied in two different fonts such that one font was used to study a minority (two) of its associates, lures tested in that font produced fewer errors than when they were tested in the font used to study the majority (eight) of its associates (Arndt, 2010). Although this outcome could be explained by differences in perceptual fluency because one font was seen twice and the other eight times, a mismatch condition indicated that the influence of simply having studied one font twice and a second font eight times had a considerably smaller effect on lure errors (approximately 5%) than when lures were tested in a font that was used to study two or eight of its associates (approximately 13%). Thus, as with studies of source attributions, studies of font-match effects for lure items suggest that lures cue retrieval of highly specific contextual information that was present when their associates were studied.

In summary, the results of studies investigating false recollection using list-learning paradigms have documented that lure errors in the DRM paradigm produce highly compelling false recollections. However, as noted throughout this chapter, subjective conviction is generally insufficient to show that lure errors are underlain by recollection-based processes. Importantly, there are several findings that suggest participants retrieve highly detailed information about the context in which a lure's associates were encoded when a lure item is tested, which in turn underlies the conviction with which they falsely recollect lures as having been studied (see Table 1, third row). These results therefore suggest that lure errors in the DRM paradigm are underlain, in part, by retrieval of authentic recollective information, and are not solely due to familiarity-based processes.

3.4. Summary of Behavioral Evidence and Implications

Taken together, evidence from investigations of misinformation effects, imagination inflation, and list-learning false memories documents numerous cases where people express strong subjective beliefs in the authenticity of false memories. However, in many cases, basic findings that have been interpreted as evidence of false recollection (e.g., believing one performed an action that was only imagined; Goff & Roediger, 1998) can be explained by familiarity-based processes because the evidence is based upon people's subjective belief they are recollecting, rather than objectively verifiable evidence that they retrieved encoding context (see Table 1, first column). Despite the viability of a familiarity-based

explanation for the subjective conviction with which people believe their false memories, clear evidence that false recollection can reflect the engagement of recollection processes has been demonstrated by the systematicity in participants' misattributions of imagined events to aspects of encoding context (e.g., Lyle & Johnson, 2006). Further, lure items in the DRM paradigm also tend to produce retrieval of specific encoding contexts that occurred during the study of their associates (Arndt, 2010; Hicks & Hancock, 2002). Both of these outcomes suggest that the false recollection people experience is accompanied by retrieval of highly specific episodic information that was encountered during the encoding of related items, which provides a memorial basis for participants' strong belief they can recollect contextual information when they experience a false memory. Thus, at least for imagination inflation and list-learning false memories in the DRM paradigm, behavioral evidence suggests that false recollection can be underlain by retrieval of authentic encoding context rather than simply being due to misattributions of strong familiarity (see Table 1, second column). Next, I review one final source of evidence that can aid understanding of the memorial bases of false recollection—studies examining false recollection with brain imaging techniques.

4. Brain-Based Evidence for False Recollection

Research examining how the brain responds to false recollections has used ERPs, PET, and fMRI to better understand the brain processes underlying false memories. Much of the research examining the brain's response to accurate and false events has sought to find differences between neural responses to events that were actually experienced and those that were not. As one would expect, authentic memories tend to produce more activation in sensory areas of the brain (Slotnick & Schacter, 2004), as well as areas of the brain associated with retrieval monitoring (Weise & Daum, 2006) compared to false memories. An equally important question for understanding the nature of memory errors is to examine similarities in how the brain responds to events that were authentically experienced and those that were not. Further, and critical for understanding the nature of false recollection, are cases where both true and false memories activate brain areas that are associated with retrieval of contextual information. Thus, in this section, I first describe the areas of the brain that are activated during retrieval of contextual information in ERPs, PET, and fMRI, and then follow that description with a review of the literature documenting cases where false memories activate those areas, with the goal of evaluating how strongly neuroimaging-based evidence supports the notion that people retrieve encoding context when they experience false recollections.

Research investigating ERPs has identified separate neural signatures that are associated with recognition memory familiarity and recollection, respectively (Rugg & Curran, 2007). Familiarity-based recognition tends to be associated with early activation (~400 ms poststimulus) that is frontally distributed. Recollection, on the other hand, is associated with later activation (~600 ms poststimulus) that is parietally distributed, and often left-lateralized. Important for the use of ERPs to understand the memory processes underlying false memories is that later parietally distributed ERP activation is sensitive to outcomes such as accurate source retrieval (Cansino & Trejo-Morales, 2008; Wilding, 1999), the ability to discriminate between study items' exact form and plurality-reversed test lures (e.g., *frog* vs. *frogs*; Curran, 2000), the ability to accurately discriminate studied image pairs from test pairs composed of two images that were studied, but as part of different image pairs (Speer & Curran, 2007), and the number of details a person is able to retrieve when they remember a study item (Vilberg, Moosavi, & Rugg, 2006). Each of these findings suggests that parietal ERPs reflect retrieval of specific details that were stored during an encoding episode (i.e., they reflect retrieval of episodic information, such as encoding context). Thus, the hypothesis that false recollection is, in part, underlain by retrieval of encoding context suggests that false recollections should produce parietal ERPs.

Research investigating basic memory processes using the neuroimaging techniques PET and fMRI has identified the medial temporal lobe (MTL) as critical for context-based memory, such as memory for the source of an experience (Mitchell & Johnson, 2009). Thus, cases in which the MTL is activated by memories for events that were not experienced provide support for the hypothesis that false recollection is underlain by some of the same basic memory processes as authentic recollection, in particular storage and retrieval of contextual information that is associated with episodic memories. Although other brain areas that are activated by false memories are also activated by true memories (e.g., areas of frontal and parietal cortex; Garoff-Eaton, Slotnick, & Schacter, 2006), the critical brain areas for assessing the hypothesis that false recollection is underlain by retrieval of encoding context is whether those false memories activate the medial temporal lobes. Thus, the review below will focus on instances where false recollection was and was not accompanied by parietal activation in ERP studies, and activation of MTL structures in PET and fMRI studies.

4.1. ERP Studies

The vast majority of the studies using ERP to investigate false memory have been conducted by examining lure errors in the DRM paradigm. Thus, although there are studies examining how imagining and

performing actions impact ERPs (Johansson, Stenberg, Lindgren, & Rosén, 2002; Leynes & Bink, 2002; Senkfor, Van Petten, & Kutas, 2002), those studies have not investigated ERPs for cases where actions that were imagined were falsely remembered as performed. Studies of lure errors in the DRM paradigm have shown parietal activation for lures in time windows that encompass 600 ms poststimulus (Johnson, Nolde, Mather, Kounios, Schacter, & Curran, 1997; Nessler & Mecklinger, 2003; Nessler, Mecklinger, & Penney, 2001; Weise & Daum, 2006), although not all ERP studies have shown such activation (Curran, Schacter, Johnson, & Spinks, 2001; Geng, Qi, Li, Fan, Wu, & Zhu, 2007). Further, the parietal activation found in this time window is not always restricted to the left parietal locations that are most strongly associated with recollection (Rugg & Curran, 2007). Instead, the activation in parietal areas is often broadly distributed across the brain (Johnson et al., 1997; Nessler et al., 2001). Thus, although DRM lure items often produce parietal ERP activation that is characteristic of recollection, it is also often the case that the activation produced by lure items is not specifically seen in parietal areas whose activation is believed to be a neural signature of recollection. As a consequence, the ERPs that lure items in the DRM paradigm produce do not strongly support the theory that those ERPs arise from recollection-based processes.

That ERPs do not support the hypothesis that lure errors arise from recollection-based processes is puzzling in light of the strong behavioral evidence, reviewed earlier, that lure errors produce retrieval of highly specific contextual information present when their associates were studied. Although it is hard to determine the exact reason for this discrepancy, one possibility is that ERP studies of false recollection have tended to use stimuli that are somewhat different than those used in behavioral studies that provide evidence that lures produce retrieval of encoding context. For example, Nessler and Mecklinger (2003) had participants study four members of a given category (e.g., birds), and used other category members as lures. In contrast, behavioral studies using the DRM paradigm have participants study lists of associates of a single word, all of which produce the lure in free association, and typically have participants study 10 or more of a lure's associates (Roediger & McDermott, 1995). Similarly, although some of the ERP-based studies of DRM lure errors used standard DRM study lists, those that did classified multiple items from each list as lures (e.g., the lure as well as other words that generally fit with the theme of a lure's associates) rather than simply the lure item that is known to be associated to all of the studied items in a given list (Wiese & Daum, 2006). Classifying items that are not necessarily associated with all study items from a DRM list as a lure item likely had the effect of reducing lure errors relative to standard conditions used in behavioral studies, and thus may have moderated the neural signals that were produced in response

to the errors that were made to items classified as lures in these studies. Thus, although ERPs do not provide strong support for the notion that lure items produce recollective signals, there are discrepancies between how behavioral research and ERP-based research were conducted that may explain the discrepancy between the outcomes of ERP-based studies and those of behaviorally based studies.

4.2. fMRI and PET Studies

While evidence from ERP-based studies does not strongly support the notion that false memories are associated with context, studies using PET and fMRI have provided stronger evidence suggesting false recollections are, in part, produced by encoding and retrieval of contextual information. For example, studies of misinformation effects have shown that misinformation that was mistakenly endorsed as being part of an originally witnessed event produced MTL activation at retrieval (Stark, Okado, & Loftus, 2010), as well as during encoding (Baym & Gonsalves, 2010). Further, one report investigated the neural signals accompanying encoding during study of the originally witnessed event as well as during the provision of misinformation (Okado & Stark, 2005). This study found that MTL activation during encoding of the originally witnessed event was greater for events that were later correctly remembered compared to events where misinformation was later falsely remembered. Critically, events that were later falsely remembered showed greater MTL activation during encoding of misinformation compared to events that were later remembered correctly. This finding suggests that there is a tendency in misinformation effect studies for items that produced greater MTL activation to be later remembered as part of the originally witnessed event, regardless of whether they were experienced as part of the witnessed event or as part of the misinformation narrative. The important implication of this finding is that although MTL activation for misinformation was greater for events that were later falsely remembered, suggesting participants should have had the context of the misinformation narrative more available to them at retrieval, people may instead either confuse the context present during the misinformation narrative with that from the originally witnessed event or they may retrieve the witnessed event as they are thinking about the misinformation, potentially integrating the misinformation in to their memory of the witnessed event (Okado & Stark, 2005).

Similar to neuroimaging studies of misinformation effects, numerous studies investigating false memories in the DRM paradigm have found that lure items that are semantically related to studied items produce MTL activation (Abe, Okuda, Suzuki, Sasaki, Matsuda, & Mori, 2008; Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Schacter, Buckner, Koutstaal,

Dale, & Rosen, 1997), although not all studies using DRM-like paradigms have found this result (Garoff-Eaton, Kensinger, & Schacter, 2007; Kim & Cabeza, 2007a; Kim & Cabeza, 2007b; von Zerssen, Mecklinger, Opitz, & von Cramon, 2001). One possibility for this discrepancy is that most of the studies that failed to show MTL activation for lure items used categorized stimuli (Kim & Cabeza, 2007a,b; von Zerssen et al., 2001) rather than stimuli that are most commonly used in DRM studies where all study items are associated with a single, unstudied lure item. Further, the study that did not use categorized stimuli (Garoff-Eaton et al., 2007) was designed to examine differences between conceptual and perceptual false recognition. Thus, stimulus construction was constrained by the need for lure items to have both associates that were conceptually related and associates that were perceptually related. This constraint meant that only a small number of associates, three, were shown for each lure item, which likely limited how strong the false memory illusion was for participants (Robinson & Roediger, 1997), and may have lowered the likelihood that lures would also produce MTL activation.

Finally, while there is not a great deal of comparable research in the behavioral literature, two fMRI-based studies have investigated false memory for lures that are perceptually similar to studied images (Garoff-Eaton et al., 2006; Slotnick & Schacter, 2004). In both of these studies, MTL activation was found for both false alarms to perceptually similar lures as well as hits to studied images. Further, and critical to the hypothesis that similar lure items are unique in showing MTL activation, Garoff-Eaton et al. (2006) documented that false alarms to perceptually similar lures produced greater MTL activation than false alarms to unrelated lure items. Thus, rather than MTL activation simply occurring whenever a person believes an item was studied, MTL activation seems to be specific to cases where a lure item is related to events that were actually experienced.

4.3. Summary of Brain-Based Evidence and Implications

Table 1 summarizes the evidence for ERP (third column) and PET/fMRI (fourth column) studies of false recollection. While ERP-based research is equivocal regarding whether or not false memories produce retrieval of encoding context, PET- and fMRI-based measures of brain activity suggests that false recollection that results from the provision of misinformation and testing DRM lure items produces retrieval of encoding context. These outcomes provide evidence to complement the outcomes of behavioral studies. Importantly, in the case of misinformation effects, where the behavioral evidence was unable to falsify a familiarity-based explanation of false recollection, the results from neuroimaging studies suggest that misinformation-based false memories activate brain areas

associated with encoding and retrieval of contextual information, which supports the notion that misinformation can produce erroneous retrieval of encoding context. Further, PET- and fMRI-based studies of lure errors in the DRM paradigm tend to show that lure errors activate the MTL, a brain area that plays a key role in encoding and retrieving contextual information. Such results underscore the behavioral evidence that lures generate retrieval of specific details of the encoding context that was present when a lure's associates were encoded. Thus, between behavioral and neuroimaging-based evidence, the false recollections that accompany the misinformation effect, imagination inflation, and DRM lure errors all have at least one line of evidence suggesting that the false memories they produce are, in part, underlain by retrieval of encoding context. Next, I turn to consider how these findings can be used to inform theoretical views of the nature of the memory processes underlying false memory.

5. Theories of False Recollection

There are three primary theories that make explicit assumptions about the representations that underlie false memories, as well as how accessing those representations produces false recollection: fuzzy-trace theory (Brainerd, Reyna, & Kneer, 1995), activation-monitoring theory (Roediger, Watson, McDermott, & Gallo, 2001), and global-matching models (Arndt & Hirshman, 1998; Hicks & Starns, 2006b; Hintzman, 1988). Each of these theories makes fundamentally different claims about the memory representations underlying false memories, which in turn leads them to produce different explanations for the occurrence of false recollection. These theories' accounts are reviewed, followed by an assessment of how well the theory accounts for the major phenomena documenting that false recollection represents authentic retrieval of encoding context. A fourth perspective that is commonly used to understand memory errors is the source monitoring framework (Johnson et al., 1993). Although this perspective also seeks to explain many of the findings outlined above, there are two primary reasons I do not consider it here. First, its account of memory errors is covered elsewhere (e.g., Johnson, 2006; Mitchell & Johnson, 2009). Second, the source monitoring framework tends to focus on the retrieval processes people use when deciding whether an event was experienced, as well as the context in which that event was experienced, but makes somewhat more general assumptions about the nature of the representations that false memories and false recollections access than the theories described in detail here. Thus, this chapter will focus on theories that describe the representations that underlie false memory in detail, with the goal of better understanding the nature of the representations that are accessed when people experience false recollection.

5.1. Fuzzy-Trace Theory

Fuzzy-trace theory's explanation of false recollection depends on the type of item that produces the false memory. For items that were not encountered prior to a memory test (e.g., DRM lures), the theory claims false memory is due to familiarity-based processes. For items that were encountered during an experiment (e.g., misinformation or imagined items), the theory proposes an explanation using recollection-based processes. Although this chapter has already reviewed evidence documenting the limitations of familiarity-based explanations of false memories, fuzzy-trace theory provides a specific implementation of that viewpoint, and does so within the framework of a memory theory that has been widely applied to false memory phenomena (e.g., Brainerd & Reyna, 2002; Brainerd, Reyna, & Ceci, 2008; Brainerd, Reyna, Wright, & Mojardin, 2003). Thus, although the theory's claim that false memories in the DRM paradigm are driven by familiarity-based processes is insufficient to explain the data for DRM lure errors reviewed in this chapter, the theory has enjoyed substantial success accounting for false memory phenomena generally. Further, in contrast to the general familiarity viewpoint described to this point, fuzzy-trace theory makes explicit assumptions about the representations underlying familiarity. Finally, for phenomena fuzzy-trace theory explains with recollection-based processes, the theory makes explicit assumptions about the nature of the representations that can and cannot produce false recollection. Thus, examining this specific theory in light of the data reviewed in this chapter can provide an illustration of the challenges faced by theories that attempt to explain false recollection in terms of familiarity-based processes, and will provide a test of one set of claims about how recollection-based processes can produce false memory.

Fuzzy-trace theory (Brainerd et al., 1995) argues that any event one experiences produces storage of memory traces that vary on a continuum between gist traces and verbatim traces. Verbatim traces represent the details of an experience, such as its perceptual characteristics, while gist traces represent the commonalities among previously encountered events, such as the semantic features they have in common. Importantly, verbatim trace retrieval underlies the ability to recollect an event's prior occurrence while gist trace retrieval underlies the experience of familiarity with a previous event. Fuzzy-trace theory argues that memory errors for unstudied events arise from those events matching gist traces. Thus, fuzzy-trace theory argues that memory errors for unstudied items arise because the unstudied items are similar to studied items along some dimension represented in a gist memory trace, such as semantics. Finally, in order to explain the subjectively compelling nature of memory errors that people experience for unstudied events, fuzzy-trace theory suggests people can

experience a phenomenon known as *phantom recollection* (Brainerd, et al., 2001), where people mistake the strong familiarity that results from a test item strongly matching a gist memory trace for the experience of recollecting. Thus, although phantom recollection gives rise to the subjective experience of recollecting, it is based upon retrieval of a gist trace, which mediates familiarity in fuzzy-trace theory.

In order to explain misinformation effects, fuzzy-trace theory proposes there are two reasons people will confuse misinformation with an event they actually witnessed. First, people may accidentally retrieve the verbatim trace that was formed when they encoded the misinformation (Brainerd & Reyna, 1998). Second, the misinformation may match the gist of their memory for the actual event (Brainerd & Reyna, 1998). Thus, although fuzzy-trace theory suggests people may mistakenly believe misinformation was part of an originally encoded event because it is familiar, the theory also suggests that people can *recollect* misinformation by retrieving the verbatim memory trace from when they were presented with it, which could, in turn, lead them to believe the misinformation was experienced during the originally witnessed event. This account, because it relies on mistaken recollection of misinformation, can explain data showing that misinformation that is later mistakenly believed to be part of the originally witnessed event produced MTL activation during exposure to misinformation (Baym & Gonsalves, 2010), during retrieval of misinformation (Stark et al., 2010), and that misinformation which was later falsely believed to be part of the originally witnessed event produced greater MTL activation during exposure to misinformation than events that were later correctly remembered (Okado & Stark, 2005).

Fuzzy-trace theory offers a similar explanation for imagination inflation as it does for misinformation effects. Thus, imagined items produce both verbatim and gist memory traces, such that false recollection could occur to the extent that imagined items produced retrieval of a verbatim trace on a later memory test. Presumably, this verbatim memory trace would contain perceptual information that was sufficiently confusable with actual perceptual information, given that imagery and perception have a number of similarities behaviorally (Kosslyn, Ball, & Reiser, 1978) and neurally (Ganis, Thompson, & Kosslyn, 2004). Although fuzzy-trace theory can generally explain why people believe imagined items were perceived, as well as that they may actually retrieve encoding context as a basis for that belief, one challenge for fuzzy-trace theory is to specify mechanisms that allow it to explain why people tend to believe imagined items were encountered in the source used to encode perceived items that are perceptually and conceptually related to imagined items (Lyle & Johnson, 2006). The idea that people mistakenly confuse imagined items' verbatim traces with perceived items' verbatim traces could be explained if the verbatim trace of an imagined item sufficiently resembled that of a

perceptually similar object that was actually perceived, which in turn would lead people to believe that the imagined item was encoded in the location or color of the perceived item. Although the exact mechanisms that would allow such a confusion are not currently specified by fuzzy-trace theory, perceptually based confusions seem plausible given that verbatim traces represent the surface form of objects that were encoded. However, the fact that imagination inflation research has documented that imagined items which are conceptually related to perceived items also produce systematic beliefs about the source in which an imagined item was encountered (Lyle & Johnson, 2006) appears to be difficult for fuzzy-trace theory to explain. In particular, semantic information is represented in gist traces in fuzzy-trace theory, which produce familiarity-based recognition. Given that gist traces do not represent the details of a prior experience, it is difficult to understand how conceptual similarity could produce a systematic bias to attribute an imagined item to the same source as a conceptually related item that was perceived. Thus, at least some of the data documenting that people retrieve specific source information when memory is tested with imagined items seems to be inconsistent with fuzzy-trace theory's explanation of false recollection.

Finally, fuzzy-trace theory experiences difficulty explaining evidence that lure errors in the DRM paradigm produce retrieval of specific encoding contexts that were used to present the lure's associates during encoding. Fuzzy-trace theory attributes lure errors in the DRM paradigm to lure items matching gist representations that were formed during the encoding of the lure's associates. Specifically, gist traces represent the semantic commonalities among a lure's studied associates, and lure items tend to strongly resemble those gist traces that were formed when the lure's associates were encoded. The theory explains why people find lures to be subjectively compelling by suggesting that strong familiarity can underlie phantom recollection, which produces the subjective experience of recollecting events that did not occur (Brainerd et al., 2001). However, given that fuzzy-trace theory argues that lure errors are caused by retrieval of gist traces, the theory claims that lure errors are underlain by familiarity-based representations. Recall that familiarity-based representations in general, and the gist-based familiarity representations in fuzzy-trace theory in particular, do not represent specific details of the experience that produced them, such as the source or context in which the trace was formed. This consideration leads fuzzy-trace theory to expect that source judgments for lure items will be made at random rather than in a way that produces systematic biases to attribute lures to the source used to present a lure's associates during encoding (Anastasi et al., 2000; Roediger et al., 2004), or to attribute the lure to the source of its strong associates more often than the source of its weak associates (Hicks & Hancock, 2002; Hicks & Starns, 2006a). Similarly, fuzzy-trace theory does not appear to

be able to explain why lure errors increase when they are tested in the font used to study their associates compared to when they are tested in a font that was studied, but was not used to study the lure's associates (Arndt, 2010), because gist-based memory representations do not contain information about the specific encoding context that was present when they were formed. Thus, fuzzy-trace theory will require modification to explain evidence that lure items in the DRM paradigm produce retrieval of encoding context.

In sum, fuzzy-trace theory is capable of accounting for people's subjective sense that they can recollect misinformation from a witnessed event, that people believe imagined items were perceived, and people's belief that DRM lures were encountered in a study list. Further, the theory is able to explain why people show evidence of retrieving encoding context for misinformation, as indexed by fMRI activation during encoding (Baym & Gonsalves, 2010; Okado & Stark, 2005) and retrieval (Stark et al., 2010). The theory encounters difficulty, however, explaining the specificity of the knowledge people have about their false recollections. In particular, it appears unable to account for the bias to attribute imagined items to the source of conceptually related perceived items (Lyle & Johnson, 2006), as well as the specific source (Anastasi et al., 2000; Hicks & Hancock, 2002; Hicks & Starns, 2006a; Roediger et al., 2004) and contextual information (Arndt, 2006, 2010) people retrieve when their memory is tested with lure items in the DRM paradigm. While some of these challenges arise due to fuzzy-trace theory's claim that unstudied items produce false recollection from familiarity-based processes (Arndt, 2010), even in cases where the theory invokes recollection-based processes to explain imagination inflation, the theory appears to encounter difficulty explaining the breadth of stimulus dimensions that produce erroneous source retrieval for imagined items. Thus, fuzzy-trace theory will require further development in order to explain all of the data that document false recollection, particularly the specificity of information people retrieve when they are tested with an item that was not perceived or studied.

5.2. Activation-Monitoring Theory

In contrast to fuzzy-trace theory, activation-monitoring theory (Roediger et al., 2001) has properties that potentially allow it to explain why false memories are accompanied by retrieval of encoding context in most situations. Specifically, activation-monitoring theory proposes that encoding-based processes can lead to the formation of associations between representations for events that did not occur and encoding context. These associations, when reactivated during retrieval, form the basis for activation-monitoring theory's ability to explain why some false

memories are accompanied by retrieval of encoding context. Importantly, although activation-monitoring theory is not constrained to claim that all false memory phenomena are largely encoding-based, its explanation of false recollection is primarily due to processes that occur during encoding. Thus, activation-monitoring theory's explanation of compelling false memories potentially provides insight into the successes and challenges faced by theories that place the locus of false recollection at encoding.

Activation-monitoring theory was developed to explain lure errors in the DRM paradigm. The theory's explanation of DRM lure errors is that during the course of encoding a lure's associates, activation spreads from the associates' representations to the lure's representation in a semantic network (Collins & Loftus, 1975). At test, lures access a highly active representation in semantic memory, which in turn provides substantial evidence the lure was studied. However, representation activation is a source of evidence that does not indicate details about the basis for the representation's activation, making it more similar to recognition memory familiarity than to the retrieval of encoding context that was argued above to underlie false recollection in the DRM paradigm. Activation-monitoring theory suggests that false recollection can arise when lures are particularly strongly activated during encoding, which can cause lures' representations to become associated with encoding context. Associations that are formed between lure representations and encoding context allows activation-monitoring theory to explain the general occurrence of false recollection by suggesting that when people are tested with lure items, they retrieve evidence that lures were, in fact, encountered in the same encoding context as studied items, producing the experience that they can recollect the lure's occurrence during encoding.

As a result of its claim that lure representations can become associated with encoding context, activation-monitoring theory can account for most of the basic findings that demonstrate that lures produce retrieval of specific details of encoding context (Arndt, 2010; Hicks & Hancock, 2002). For example, it can explain why people believe lures were more likely to have been studied in the source used to present the lure's strong associates (Hicks & Hancock, 2002; Hicks & Starns, 2006a) by suggesting that the lure's strong associates activate the lure more during encoding, which causes the lure's representation to be more strongly associated with the encoding context (source) present when the lure's strong associates were encoded compared to the encoding context (source) present when the lure's weaker associates were encoded. Similarly, activation-monitoring theory explains why people make more mistakes to lures that were tested in the same font that was used to study its associates compared to a studied font that was not used to study its associates (Arndt, 2010). In particular, the theory suggests that lures become associated with the encoding context that was present when it was activated by its studied

associates. Thus, the features of encoding context that were present at that point in time, such as a font that was used to present the associates, should be linked to the lure's representation in semantic memory. Importantly, when that association is reinstated at retrieval, it will inflate lure errors relative to when the association is not reinstated at retrieval, as occurs when lures are tested in a font that was not used to encode its associates. Finally, because activation-monitoring theory suggests that false recollection occurs because lure representations become directly associated with encoding context, the theory provides a straightforward explanation for why lure items often produce brain activation in areas that are associated with encoding and retrieval of contextual information (e.g., MTL activation). In sum, activation-monitoring theory provides an explanation for most aspects of false recollection in the DRM paradigm.

Although activation-monitoring theory was not designed to explain misinformation effects and imagination inflation effects, an explanation can be derived from its general principles. The general occurrence of misinformation effects can be explained by activation-monitoring theory as a consequence of the fact that misinformation was experienced during a study, causing its representation in semantic memory to be active. This activation will give a person evidence that the misinformation was experienced recently, which could be confused with the experience of encoding misinformation as part of the originally witnessed event. Further, cases where misinformation produces MTL activation can be thought of as a reflection of the encoding and retrieval of associations between the misinformation's representation in semantic memory and a representation of the context that was present during the provision of misinformation. Thus, MTL activation at encoding can be explained as reflecting storage of context for misinformation items, and MTL activation at retrieval could be a reflection of confusion between the context in which the misinformation occurred and the context of the originally witnessed event (see Ayers & Reder, 1998 for a similar explanation). A second explanation activation-monitoring theory could advance for why misinformation produces MTL activation at both encoding and retrieval is that when misinformation is provided, it activates the originally witnessed event's representation, which in turn enables the misinformation to become associated with the context of the originally witnessed event. This explanation would readily account for the occurrence of MTL activation at both encoding and at retrieval when misinformation is confused with an originally witnessed event, because the misinformation was actually associated with the encoding context of the originally witnessed event. In either case, the basic principles of semantic memory activation and association of semantic memory representations with encoding context advocated by activation-monitoring theory seem to be capable of explaining false recollections that result from the provision of misinformation.

Activation-monitoring theory would advance a similar explanation for imagination inflation as it does for misinformation effects. In particular, the theory would assume that perceived/performed items and imagined items both activate their respective representations in semantic memory. In order to explain the general finding of imagination inflation, activation-monitoring theory could claim that people confuse the activation of representations for imagined items with that from performed/perceived items. In order to explain evidence that people retrieve specific details of encoding context when their memory for imagined items is tested (Lyle & Johnson, 2006), activation-monitoring theory would require additional assumptions regarding how context is associated with perceived and imagined items. Similar to its explanation of lure errors in the DRM paradigm, activation-monitoring theory would expect that semantically related representations are activated when an object is perceived. This activation from the perceived object's representation in semantic memory would tend to activate the semantic representation for imagined items that are conceptually related to the perceived item. Activation of imagined items' representation in semantic memory would have the potential to cause those imagined items' representations to become associated with encoding context that was present when a perceived item was studied. While this allows the theory to explain cases where imagined items are attributed to the source of conceptually similar perceived items, activation-monitoring theory does not appear to have a straightforward mechanism that would allow it to explain why an imagined item tends to be attributed to the same source as a perceived item that was *perceptually* similar to it. That is, because activation-monitoring theory claims that imagined items are associated with related items' context when they receive activation from the perceived items' representations in semantic memory, it does not have a mechanism that enables perceptual similarity to activate an imagined item's representation in semantic memory, and thus does not have a mechanism that can lead to the formation of an association between an imagined item and the source characteristics of a perceptually related item that was perceived.

In summary, while activation-monitoring theory can explain much of the data documenting false recollection, it does not appear to provide a comprehensive account of all of the stimulus dimensions that can produce retrieval of encoding context when a false recollection is experienced. Importantly, the challenges faced by the theory appear to be due to its explanation that false recollection is caused by associating false memory items (e.g., imagined items) with contextual information (e.g., that the item was perceived in a particular place or color) during encoding. While those associations must also be cued at retrieval, the primary process that enables their retrieval is that the association between a false memory item and encoding context was constructed during encoding—a process that

requires the false memory item's representation to be activated during encoding. Thus, when activation of a false memory item's representation in semantic memory is unlikely to occur (e.g., when no semantically related items are encountered during encoding), activation-monitoring theory does not have a mechanism that allows it to explain false recollection. Importantly, the challenges activation-monitoring faces in explaining some aspects of false recollection are illustrative of the challenges faced by explanations that rest heavily on encoding-based processes—such theories must provide a mechanism by which unstudied (or unperceived) information is activated during encoding, which in turn produces associations with encoding context.

5.3. Global-Matching Models

Global-matching models (Arndt & Hirshman, 1998; Hintzman, 1988) take a very different approach to explaining memory errors compared to activation-monitoring theory and fuzzy-trace theory. Specifically, global-matching models propose that during encoding, each event that is experienced produces a memory trace that contains features describing the item's semantics, perceptual characteristics, and the context in which it occurred. During later memory retrieval, test items are compared to all of the traces in memory, producing an activation value based upon the similarity of each trace to the tested item. Those similarity values are then summed, producing an overall value for how much memory is activated by the test item. If the overall activation of memory is sufficiently high, the test item will be judged to be studied. As applied to lure errors in the DRM paradigm, global-matching models claim that memory errors arise because lures somewhat match the memory traces of their studied associates. Thus, although lure items may not bear strong resemblance to any individual trace in memory, they can accrue substantial evidence they were studied based upon the fact that they will activate a number of memory traces to a moderate degree, because the summation of those moderate activations will produce relatively strong activation of memory as a whole.

The general characterization of memory errors offered by global-matching models, as well as the application of these models to false memory phenomena, demonstrates that global-matching models argue that false memory primarily arises during retrieval. Although the models do make claims about the nature of encoding, the models' explanation of false memory phenomena arises from how their global-matching mechanism interacts with memory traces. Further, the encoding-based processes in global-matching models simply rely on storing memory traces that reflect the stimulus information available during encoding. Thus, in contrast to activation-monitoring theory, global-matching models

suggest that compelling false memories primarily arise from the nature of retrieval processes, rather than how representations are activated and associated with context during encoding. Further, as will be detailed in the following paragraphs, global-matching models explain why compelling false memories are sometimes accompanied by retrieval of encoding context, and do so using retrieval-based mechanisms. Thus, they also differ from fuzzy-trace theory in terms of the memorial information they argue underlies compelling false memories for unstudied items, enabling them to provide a potential explanation for all of the false recollection phenomena reviewed above.

In order to explain retrieval of encoding context, global-matching models use the same basic encoding and retrieval mechanisms described above, and simply assume that studied items' memory traces are composed of both information about the items that were encoded, as well as elements of the context in which they were experienced (e.g., the voice or font in which it was encoded). In order to make source memory judgments, global-matching models assume participants construct a separate memory probe for each source that is composed of the item being tested and each of the sources the person is being asked to judge. They then probe memory with a representation that combines the test item and each source. The probe using the source that produces the greatest activation of memory determines which source will be judged as the source of that test item. Critical for their ability to explain source judgments for studied items is that global-matching models possess a characteristic known as *interactive cueing* (Clark & Gronlund, 1996). Interactive cueing is the property that when a test probe composed of multiple types of information (e.g., item information and source information) matches a memory trace in terms of both its item information and its source information, it generates a larger activation value than when a memory probe matches item information and source information that is contained in two different memory traces. Thus, when a test probe matches both item and source information in the same memory trace, that test probe will tend to, on average, produce a higher overall activation of memory than when the test probe matches item and source information in different memory traces (Hicks & Starns, 2006b).

As applied to explaining lure source judgments, global-matching models advance a similar explanation (Hicks & Starns, 2006b). Specifically, in order to make a source judgment for lure items, a test probe is constructed that contains item information from the lure item and source information from each of the sources a person is asked to judge. Thus, when a lure's associates were all studied in a single source, the lure's test probe that was constructed with that source will tend to match item information in a number of memory traces (those of its studied associates) as well as source information in those same traces. In contrast,

when the lure's test probe was constructed with a source that was not used to study its associates, it will match the item information in the memory traces that correspond to its studied associates, but will only match source information in memory traces where it does not tend to also match item information. As a result, interactive cueing will tend to happen for when a lure's test probe contains the source of its studied associates, but not when a lure's test probe contains a source other than the one that was used to study its associates. Thus, a lure's test probe will tend to produce greater overall activation of memory when the probe contains source information that was encoded with its studied associates compared to when the probe contains source information that was not encoded with its associates, producing a bias to attribute lures to the same source as their studied associates, the exact result that is typically found (Anastasi et al., 2000; Hicks & Hancock, 2002; Hicks & Starns, 2006; Mather et al., 1997; Roediger et al., 2004). Further, global-matching models can explain why lures tend to be attributed to the source used to encode their strong associates when the lure's associates are studied in two different sources (Hicks & Hancock, 2002; Hicks & Starns, 2006a). Specifically, global-matching models assume that lures are more similar to their strong associates than with their weak associates. Thus, lure items will produce a larger match to item information for its strong associates than its weak associates. While the lure will produce interactive cueing for both the test probe corresponding to the source of lure's strong associates and the test probe corresponding to the source of the lure's weak associates, the fact that the lure will generate a larger match to the strong associates' item information means it will also produce a greater degree of interactive cueing when it matches both item and source information for the lure's strong associates than when it matches both item and source information for the lure's weak associates. Thus, global-matching models can also explain why people show a bias to attribute lures to the source of their strong associates (Hicks & Hancock, 2002; Hicks & Starns, 2006a).

Global-matching models use the same basic principles to explain font-match effects on lure errors (Arndt, 2010). Thus, lures tested in a font that matches the font used to encode its associates produce a memory probe that contains both item and source (contextual) information that matches the same representations in memory—those of the lure's associates which were encoded in that font. In contrast, lures tested in a font that does not match the font used to encode its associates produce a memory probe that contains item and contextual information that matches different representations in memory, such that item information matches the item information from the lure's studied associates, and the contextual information in the memory probe matches the contextual information for the items that were encoded in the font the lure is tested in. Thus, interactive

cueing will occur in the match condition, but not in the mismatch condition, producing greater levels of activation for lures in the match condition compared to the mismatch condition, and consequently higher levels of lure errors in the match condition. Similarly, global-matching models explain the fact that lure errors are higher when lures are tested in a font used to study the lure's strong associates compared to when they are tested in a font used to study the lure's weak associates (Arndt, 2006) using the same basic explanation as they do to explain people's bias to attribute lures to the source used to present its strong associates during encoding. In particular, lures will share greater similarity with their strong associates' memory representations than their weak associates' memory representations, which in turn produces a stronger overall memory signal when a lure is tested in the font (context) used to study its strong associates compared to when it is tested in the font used to study its weak associates. In sum, global-matching models account for all of the major findings in the DRM literature that support the notion that lures produce retrieval of encoding context, and do so using the same basic principles they use to explain recognition and source memory for studied items, such that they provide a highly parsimonious explanation for false recollection in the DRM paradigm.

Global-matching models offer a similar explanation for why imagination produces retrieval of encoding context of items that were perceived (Lyle & Johnson, 2006). In order to explain imagination inflation, global-matching models claim that imagination inflation studies produce a series of memory traces that represent information about the items that were experienced during an experiment, as well as other details of their presentation, such as the mental operations that were undertaken when they were presented (e.g., if it was perceived, if an action was executed, or if the item was imagined). At test, people construct a memory probe using both the test item, as well as the sources they were asked to judge (e.g., perceived/performed vs. imagined). Global-matching models would assume that imagination inflation generally occurs because of the similarity between imagination and action or perception. More important for testing theories' explanation of imagination inflation is whether the theory can explain (1) why items that were imagined and are perceptually or conceptually similar to a perceived picture are more likely to be judged as perceived than imagined items that are not similar to a perceived picture (Henkel & Franklin, 1998), and (2) why people show a bias to choose the source (location or color) of a perceived picture that is perceptually or conceptually similar to an imagined item when they falsely believe that imagined item was perceived (Lyle & Johnson, 2006).

Global-matching models' explanation for why imagining items that are similar to perceived items increases imagination inflation is based upon the similarity-based matching processes that are at the heart of their

memory retrieval mechanisms. Specifically, memory probes for imagined items that are similar to items that were perceived or performed will have a tendency to match both the memory trace that was created by imagining the item and the memory trace that was created by perceiving a similar item. The fact that they will tend to match two memory traces will increase the overall activation of memory for imagined items that are similar to perceived items compared to imagined items that are not similar to perceived items, which will only match a single memory trace—the one created via imagination. Importantly, because imagined items that are similar to perceived items will somewhat resemble perceived items' traces, they will also produce an increased tendency to match that memory trace in terms of both item information and source (i.e., perceived vs. imagined) information when people are trying to determine if a test item was imagined or perceived. Thus, those imagined items that are similar to perceived items will be more likely to be judged as perceived than imagined items that are not similar to a perceived item (Henkel & Franklin, 1998). Global-matching models offer a similar explanation for why people show systematic biases to attribute imagined items to the location or color in which a related perceived item was experienced (Lyle & Johnson, 2006). Specifically, in order to arrive at source judgments, people are assumed to construct memory probes containing both item information from an imagined item and a representation of each of the sources they are asked to decide between (e.g., locations or colors; Lyle & Johnson, 2006). The imagined item's memory probe that will tend to match a perceived item's representation the best will be the probe that (1) is for an imagined item that is similar to a perceived item and (2) contains source information that matches the source in which the perceived item was encoded. Thus, this greater activation will lead people to systematically choose the source in which a similar perceived item was studied when an imagined item is judged to have been perceived, producing the biased source attributions found in studies of imagination inflation.

Finally, global-matching models can explain misinformation effects in a manner similar to how they explain imagination inflation effects. During encoding, global-matching models would propose that memory traces are formed for both the events that occurred during the witnessing of the original event, as well as for the events that were presented as part of the misinformation narrative. Further, misinformation is typically similar to events that occurred during the witnessed event, and misinformation is typically conveyed in a way that it references events that occurred during the witnessed event. This should produce memory representations for misinformation that are exceedingly similar to the representations formed for the originally witnessed event. Thus, when given a misinformation item on a later memory test, it is likely to match both the memory trace for the misinformation in memory as well as many aspects of the

memory trace for the similar item that occurred during the witnessed event. This activation created by misinformation items at test provides memorial evidence that misinformation was encountered during the originally witnessed event. Further, the fact that misinformation items potentially match memories from the originally witnessed event well enough to produce source misattributions of misinformation to the originally witnessed event suggests that global-matching models can explain neuroimaging findings documenting that misinformation items produce MTL activation during retrieval (Stark et al., 2010).

In summary, global-matching models are capable of explaining most of the major phenomena that document instances of false recollection. Importantly, global-matching models use the same basic memory representations and retrieval mechanisms to explain all of the major false recollection phenomena that were reviewed here. That is, the simple consideration of how similar items that are falsely recollected are to items that were actually studied, perceived, or were part of a witnessed event enables global-matching models to account for the basic findings demonstrating false recollection, as well as the specificity of knowledge people have about false recollections. Thus, not only do global-matching models appear to provide a comprehensive account of false recollection phenomena, they also provide that account in a parsimonious manner, by employing the same mechanisms they use to explain memory for events that were actually experienced, as well as the context in which those memories were encountered. The success global-matching models have explaining cases of false recollection seems to be representative of that experienced by a theory that places the locus of false recollection effects at retrieval. Specifically, global-matching models, as well as other retrieval-based explanations of false recollection, should be able to explain false recollection phenomena as long as a memory probe for a false memory item matches the memory trace for one or more study items and their encoded contextual information.

5.4. Comparison of Theories

Comparing the success of the accounts of the three theories considered here suggests that all three provide an adequate account of misinformation effects, although the data may not be as fine-grained and capable of distinguishing among the three viewpoints' claims as the data from imagination inflation and DRM lure errors. Further, the misinformation effect is the case where the theories' accounts appear to be the most similar, which makes misinformation effect research less likely to distinguish among them. In explaining imagination inflation, global-matching models seem to provide the most comprehensive account of the specificity with which imagined items are attributed to specific sources of related

items that were perceived. In contrast, both activation-monitoring theory and fuzzy-trace theory appear to need further development to explain at least some aspects of the specificity with which people believe imagined items were encountered in a specific source. In particular, activation-monitoring theory encounters difficulty explaining how perceptual similarity between imagined and perceived items can produce source judgments for imagined items, while fuzzy-trace theory encounters difficulty explaining how conceptual similarity between imagined and perceived items can produce source judgments for imagined items. Finally, regarding DRM lure errors, both activation-monitoring theory and global-matching models generally explain the specificity of the context people retrieve when they are tested with lure items, although global-matching models seem to do slightly better in the details (Arndt, 2010). In contrast, fuzzy-trace theory appears to need further development to explain the specificity of the contextual information that is retrieved when lure items are tested, which is likely a reflection of the theory using familiarity-based representations to explain false recollection for items that were not encountered in an experiment prior to test. Thus, in sum, global-matching models appear to provide the most comprehensive explanation of false *recollection*—the highly compelling false memories that people experience in paradigms such as those used to study misinformation effects, imagination inflation, and errors in the DRM paradigm.

6. Future Considerations

In closing, there are three primary directions that are particularly important for future research to pursue in order to further our understanding of the nature of false recollection. First, understanding false recollection will be aided by additional exploration of the kinds of contextual information people are able to retrieve when they experience a false recollection. As noted throughout this chapter, the key evidence that false recollection is mediated by authentic recollection-based memory processes and representations is that false memories are sometimes accompanied by retrieval of specific episodic details. Thus, future studies would benefit from further examination of the types of contextual features that are retrieved when a false recollection occurs, as well as exploration of the kinds of memory representations and processes that enable retrieval of those contextual features.

Second, understanding false memories' bases will benefit from additional research examining the neural underpinnings of false recollection. For example, to date there are no studies of false memories resulting from imagination inflation, such that neuroimaging could produce converging

evidence that imagination can produce retrieval of encoding context indicating an imagined event was perceived. Further, research examining imagination inflation may be able to profitably explore the extent to which perceptual information distinguishes between retrieval of actually experienced and imagined events, similar to how research on imagery has documented similarities and differences between perception and imagery (Ganis et al., 2004). Similarly, understanding of false recollection may be advanced by studies examining the brain's responses to conditions that have demonstrated that encoding context is retrieved when false memories occur (Hicks & Hancock, 2002; Lyle & Johnson, 2006), and by studying conditions that have been shown to increase people's belief that they can recollect the occurrence of events that were not actually experienced, such as increasing the number of times an action is imagined (Goff & Roediger, 1998).

Third, and finally, additional studies that are motivated by testing theories of false recollection will produce two primary benefits. First, they will provide a more complete understanding of the memory representations and retrieval mechanisms that underlie false memories, particularly the false recollections that people find subjectively convincing. Second, understanding the memory representations and retrieval mechanisms underlying false recollection has potential applied benefits. Specifically, understanding the bases of false memories will enable the development of better techniques for limiting false memories as well as techniques that will better enable people to distinguish between authentic and inauthentic memories of previous events. Thus, although much has been learned about the nature of false memories and false recollections, there are many important outstanding questions that will both further theoretical understanding of human memory, and have the potential to shape applied practices in important ways.

REFERENCES

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., & Mori, E., et al., (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex*, *18*, 2811–2819.

Anastasi, J. S., Rhodes, M. G., & Burns, M. C. (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology*, *113*, 1–26.

Arndt, J. (2006). Distinctive information and false recognition: The contribution of encoding and retrieval factors. *Journal of Memory and Language*, *54*, 113–130.

Arndt, J. (2010). The role of memory activation in creating false memories for encoding context. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *36*, 66–79.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations of a global matching perspective. *Journal of Memory and Language*, *39*, 371–391.

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin and Review*, *5*, 1–21.

Baym, C. L., & Gonsalves, B. D. (2010). Comparison of neural activity that leads to true memories, false memories, and forgetting: An fMRI study of the misinformation effect. *Cognitive, Affective and Behavioral Neuroscience*, *10*, 339–348.

Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, *20*, 119–120.

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy trace theory and false memory. *Current Directions in Psychological Science*, *11*, 164–169.

Brainerd, C. J., & Reyna, V. F. (2005). The science of false memory. New York: Oxford University Press.

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, *134*, 343–382.

Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language*, *34*, 157–185.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110*, 762–784.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 307–327.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related fMRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Science*, *98*, 4805–4810.

Cansino, S., & Trejo-Morales, P. (2008). Neurophysiology of successful encoding and retrieval of source memory. *Cognitive, Affective and Behavioral Neuroscience*, *8*, 85–98.

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, *20*, 66–73.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review*, *3*, 37–60.

Collins, A., & Loftus, E. (1975). A spreading-activation theory of semantic memory. *Psychological Review*, *82*, 407–428.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory and Cognition*, *28*, 923–938.

Curran, T., Schacter, D. L., Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience*, *13*, 201–216.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Dewhurst, S. A. (2001). Category repetition and false recognition: Effects of instance frequency and category size. *Journal of Memory and Language*, *44*, 153–167.

Diana, R., Peterson, M. J., & Reder, L. M. (2004). The role of spurious feature familiarity in recognition memory. *Psychonomic Bulletin and Review*, *11*(1), 150–156.

Diana, R., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin and Review*, *13*, 1–21.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition*, *24*, 523–533.

Gallo, D. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York: Psychology Press.
Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory and Cognition*, *38*, 833–848.
Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin and Review*, *4*, 271–276.
Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, *20*, 226–241.
Garoff-Eaton, R. J., Kensinger, E. A., & Schacter, D. L. (2007). The neural correlates of conceptual and perceptual false recognition. *Learning and Memory*, *14*, 684–692.
Garoff-Eaton, R. J., Slotnick, S. D., & Schacter, D. L. (2006). Not all false memories are created equal: The neural basis of false recognition. *Cerebral Cortex*, *16*, 1645–1652.
Geng, H., Qi, Y., Li, Y., Fan, S., Wu, Y., & Zhu, Y. (2007). Neurophysiological correlates of memory illusion in both encoding and retrieval phases. *Brain Research*, *1136*, 154–168.
Goff, L. M., & Roediger, H. L., III. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recognition. *Memory and Cognition*, *26*, 20–33.
Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.
Henkel, L. A., & Franklin, N. (1998). Reality monitoring of physically similar and conceptually related objects. *Memory and Cognition*, *26*, 659–673.
Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General*, *127*, 251–268.
Hicks, J. L., & Hancock, T. W. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin and Review*, *9*, 807–815.
Hicks, J. L., & Marsh, R. L. (2001). False recognition occurs more frequently during source identification than during old–new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 375–383.
Hicks, J. L., & Starns, J. J. (2006a). The roles of associative strength and source memorability in the contextualization of false memory. *Journal of Memory and Language*, *54*, 39–54.
Hicks, J. L., & Starns, J. J. (2006b). Remembering source evidence from associatively related items: Explanations from a global matching model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1164–1173.
Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.
Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory. *Memory and Cognition*, *25*, 345–351.
Israel, I., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin and Review*, *4*, 577–581.
Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126–135.
Johansson, M., Stenberg, G., Lindgren, M., & Rosen, I. (2002). Memory for perceived and imagined pictures—an event-related potential study. *Neuropsychologia*, *40*, 986–1002.
Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*, 760–771.
Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Johnson, M. K., Nolde, S. F., Mather, M., Kounios, J., Schacter, D. L., & Curran, T. (1997). The similarity of brain activity associated with true and false recognition memory depends on test format. *Psychological Science*, *8*, 250–257.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.

Kim, H., & Cabeza, R. (2007a). Differential contributions of prefrontal, medial temporal, and sensory-perceptual regions to true and false memory formation. *Cerebral Cortex*, *17*, 2143–2150.

Kim, H., & Cabeza, R. (2007b). Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. *The Journal of Neuroscience*, *27*, 12190–12197.

Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review, 114*, 954–993.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 47–60.

Leynes, P., & Bink, M. (2002). Did I do that? An ERP study of memory for performed and planned actions. *International Journal of Psychophysiology*, *45*, 197–210.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560–572.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.

Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory, 14*, 197–213.

Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory and Cognition*, *25*, 826–837.

McCabe, D. P., & Geraci, L. D. (2009). The role of extralist associations in false remembering: A source misattribution account. *Memory and Cognition*, *37*, 130–142.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, *114*, 3–18.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212–230.

McDermott, K. B., & Roediger, H. L., III. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, *39*, 508–520.

Meade, M. L., & Roediger, H. L., III. (2002). Explorations in the social contagion of memory. *Memory and Cognition*, *30*, 995–1009.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*, 638–677.

Mitchell, K. J., & Zaragoza, M. S. (2001). Contextual overlap and eyewitness memory. *Memory and Cognition*, *29*, 616–626.

Multhaup, K. S., & Conner, C. A. (2002). The effects of considering non-list sources on the Deese–Roediger–McDermott memory illusion. *Journal of Memory and Language*, *47*, 214–228.

Nessler, D., & Mecklinger, A. (2003). ERP correlates of true and false recognition after different retential delays: Stimulus- and response-related processes. *Psychophysiology*, *40*, 146–159.

Nessler, D., Mecklinger, A., & Penney, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research*, *10*, 283–301.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition*, *25*, 838–848.

Okado, Y., & Stark, C. E. L. (2005). Neural activity during encoding predicts true and false memories created by misinformation. *Learning and Memory*, *12*, 3–11.

Payne, D. G., Elie, C. J., Blackwell, H. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, *35*, 261–285.

Reder, L. M., Donavos, D. K., & Erickson, M. A. (2002). Perceptual match effects in direct tests of memory: The role of contextual fan. *Memory and Cognition*, *30*, 312–323.

Robinson, K. J., & Roediger, H. L., III. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.

Roediger, H. L., McDermott, K. B., Pisoni, D. P., & Gallo, D. A. (2004). Illusory recollection of voices. *Memory*, *12*, 586–602.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin and Review*, *8*, 385–407.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Science*, *11*, 251–257.

Seamon, J. G., Philbin, M. M., & Harrison, L. G. (2006). Do you remember proposing marriage to the Pepsi machine? False recollections from a campus walk. *Psychonomic Bulletin and Review*, *13*, 752–756.

Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A., & Rosen, B. (1997). Late onset of anterior prefrontal activity during true and false recognition: An event-related fMRI study. *NeuroImage*, *6*, 259–269.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40*, 1–24.

Senkfor, A. J., Van Petten, C., & Kutas, M. (2002). Episodic action for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neuroscience*, *14*, 402–419.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, 7, 664–672.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 460–471.

Speer, N. K., & Curran, T. (2007). ERP correlates of familiarity and recollection processes in visual associative recognition. *Brain Research*, *1174*, 97–109.

Stark, C. E. L., Okado, Y., & Loftus, E. F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning and Memory*, *17*, 485–488.

Thomas, A. K., Bulevich, J. B., & Loftus, E. F. (2003). Exploring the role of repetition and sensory elaboration in the imagination inflation effect. *Memory and Cognition*, *31*, 630–640.

Thomas, A. K., & Loftus, E. F. (2002). Creating bizarre false memories through imagination. *Memory and Cognition*, *30*, 423–431.

Vilberg, K. L., Moosavi, R. F., & Rugg, M. D. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research*, *1122*, 161–170.

von Zerssen, G. C., Mecklinger, A., Opitz, B., & von Cramon, D. Y. (2001). Conscious recollection and illusory recognition: An event-related fMRI study. *EuropeanJournal of Neuroscience*, *13*, 2148–2156.

Weise, H., & Daum, I. (2006). Frontal positivity discriminates true from false recognition. *Brain Research*, *1075*, 183–192.

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55*, 581–598.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). On the selection of distractors for eyewitness lineups. *Journal of, Applied Psychology*, *78*, 835–844.

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia*, *37*, 441–454.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.

Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 934–945.

Zaragoza, M. S., & Mitchell, J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychological Science*, 7, 294–300.

Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memory. *Psychological Science*, *12*, 473–477.